The circadian cycle

Merrow, Martha; Roenneberg, Till

Published in:
Trends in Genetics

DOI:
10.1016/S0168-9525(00)02158-2

Publication date:
2001

The circadian cycle: is the whole greater than the sum of its parts?

Martha Merrow and Till Roenneberg

The term 'circadian rhythm' describes an oscillatory behavior in the absence of exogenous environmental cues, with a period of about a day. As yet, we don't fully understand which biological mechanisms join together to supply a stable and self-sustained oscillation with such a long period. By chipping away at the molecular mechanism with genetic approaches, some common features are emerging. In combining molecular analyses and physiological experiments, those features that are crucial for structuring a circadian day could be uncovered.

Circadian clocks are some of the miracles of Nature. Take a system that has never had constant conditions throughout evolution, remove it to an artificial, non-rhythmic environment and it continues with ~24-hour behavior, almost as if nothing has changed. Circadian rhythms are among the most common of behavioral traits, having been characterized in all phyla, regulating important and diverse functions such as photosynthesis, reproduction and locomotor activity. We have a basic understanding of the 'why?' of these rhythms: anticipation of the temporal features of the environment confers a reproductive advantage[1], and the circadian machinery equips the organism with the tools to measure time in this domain. But the 'how?' is still not fully understood. Two recent publications add new insights to the ongoing discussion. Lee *et al.*[2] describe a key post-transcriptional regulatory process in *Neurospora* involving a delay of approximately 8 h. Using a *Drosophila* long-period *doubletime* (*dbt*) mutant, Suri *et al.*[3] show that the RNA and protein oscillations, usually separated by a time-lag in circadian transcription–translation feedback loops, can coincide.

Genetic approaches have demonstrated the importance of a transcription–translation feedback loop for a robust, intact circadian system. The minimal model postulates transcription of a gene, followed by production of the protein and subsequent negative feedback on self-transcription (Fig. 1a; see Ref. 4 for the initial, basic hypothesis). Degradation of the protein then releases the negative feedback, allowing a new round of transcription and resulting in molecular oscillations of RNA and protein. Given that an artificially constructed loop can be completed in 2–3 h (Ref. 5), the circadian timescale remains an enigma. Stable insertions of lag times could slow the loops into the circadian range, and, to that end, almost any process that affects a feedback loop component is a candidate for decelerating the loop.

Regulation of expression level within the feedback loop

A key tool for circadian biologists is the experiment in constant conditions. All clock

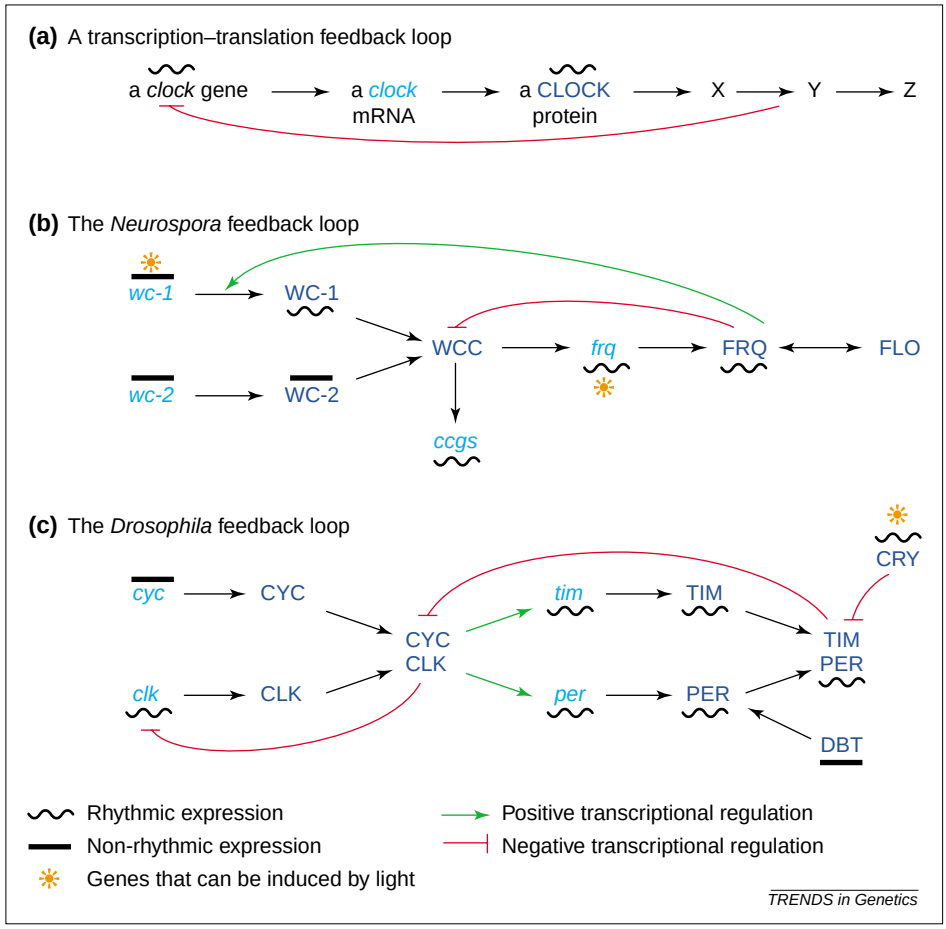

Fig. 1. Negative feedback loops in circadian systems. (a) The basic model for how clock genes function was proposed a decade ago[4]. A clock gene is transcribed, protein is translated and modified to X,Y or Z states, which permits specific information transfer, including negative feedback on self-transcription. Degradation of the clock protein allows re-initiation of transcription, thus fostering molecular oscillations of clock gene products (see Fig. 2). (b) The *Neurospora crassa* transcription–translation feedback loop[2]. Levels of *frequency (frq)* RNA and FRQ protein depend on WHITE COLLAR–1 (WC-1) and WC-2, which heterodimerize to form the WHITE COLLAR complex (WCC). WC-1 levels depend on FRQ. In constant darkness, expression of FRQ protein results in decreased *frq* RNA accumulation. The regulatory steps are inferred from genetic experiments and have not been shown directly (for instance, WCC regulates *frq* RNA but it is not clear how this is accomplished). The net effect is two interlinked regulatory loops. Light reaches the system through the WCC, which is essential for light responses in *Neurospora*. *ccg*s are clock-controlled genes, some of which are light induced. There are indications of a cryptic oscillator in the absence of this feedback loop (the FRQ-less oscillator, FLO)[18–20]. Although the FLO and the light-sensitive loop interact, it is not clear how this occurs. (c) The *Drosophila melanogaster* transcription–translation feedback loop[12]. CLOCK (CLK) and CYCLE (CYC) are factors that activate transcription of *period* (*per*) and *timeless* (*tim*) transcription. PER and TIM proteins feed back negatively on their own transcription by interfering with this process. CLK and CYC, in turn, negatively regulate *clk* RNA expression, resulting in two interlinked regulatory loops. Light reaches the system through CRYPTOCHROME (CRY)[21]. DOUBLETIME (DBT) has profound effects on the system, perhaps by phosphorylating PER directly[14].

 0168-9525/01/$ – see front matter PII: S0168-9525(00)02158-2

components identified genetically are evaluated for appropriate rhythmic expression in these 'free-running' conditions. Many clock gene RNA levels are rhythmic in a free run, suggesting transcriptional regulation. This mechanism of regulation has been demonstrated for *period* (*per*) by nuclear run-on experiments[6], and other experiments show that key components of the circadian transcription–translation feedback loops in mouse and *Drosophila* are circadianly regulated transcription factors[7]. Rhythmic protein levels could derive directly from a rhythmic RNA transcript. However, the kinetics of *per* RNA and protein accumulation indicate that post-transcriptional control mechanisms are involved on the production side[6], in addition to time-of-day-specific protein degradation[8]. The regulation of translation initiation might control the clock protein FREQUENCY (FRQ) in *Neurospora*, as is indicated by experiments using differential expression of long and short forms of the protein depending on ambient temperature[9]. Sequence analysis indicates that the *Drosophila* TIMELESS (TIM) protein could also be regulated at the level of translation initiation[10].

All of these examples pertain to proteins derived from a rhythmically expressed transcript. Lee *et al.*[2] describe post-transcriptional control involving a loop component with constitutive RNA levels in constant conditions. In the *Neurospora* feedback loop, WHITE COLLAR-1 (WC-1) and WC-2 (essential light-input pathway components) are required as positive elements for maintaining *frq* RNA and protein levels[11] (Fig. 1b). FRQ is a negative regulator of *frq* transcription. WC-1 protein levels are rhythmic in free-running conditions, although *wc-1* RNA levels are non-rhythmic. Furthermore, the overall levels of WC-1 protein are depressed in the absence of FRQ (Ref. 2). Thus, FRQ acts negatively on its own transcription and acts positively on its activator WC-1, a functional parallel with the situation in *Drosophila*, where the negative element PER has a net positive effect on the transcription of its activator, *clk*; see Fig. 1c; Ref. 12). When FRQ is artificially induced in mutants lacking the endogenous *frq* gene, WC-1 protein accumulates without a preceding increase in *wc-1* RNA. Finally, the degradation rate of WC-1 protein is independent of FRQ, making it probable that regulation of rhythmic WC-1 expression occurs at the translational level.

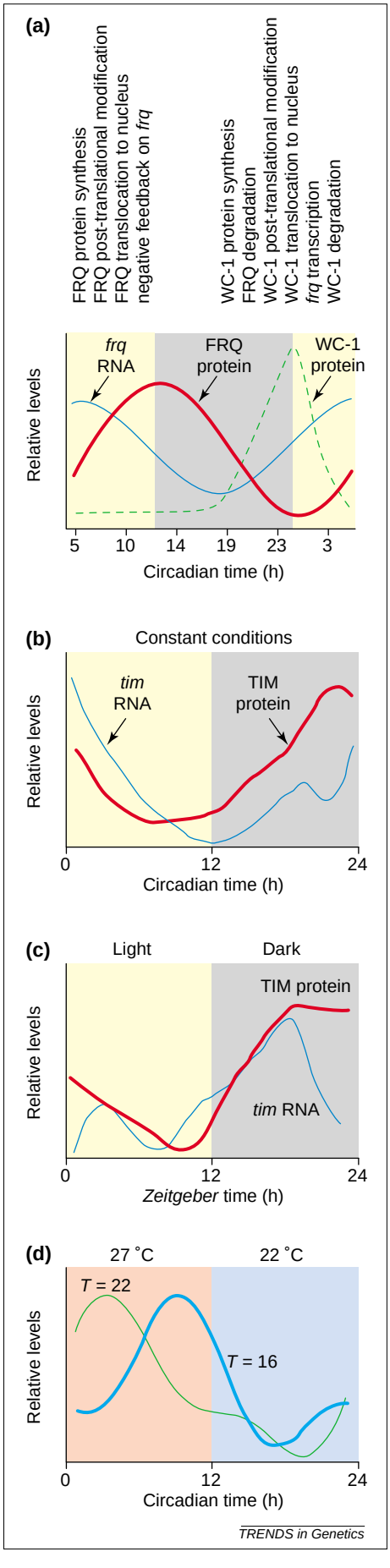

Fig. 2. Molecular oscillations of clock components. (a) The relative amounts of *frequency* (*frq*) RNA, FRQ protein and WHITE COLLAR-1 (WC-1) protein are plotted over a 22-h period, one complete 'circadian' day for *Neurospora* in constant darkness (the real hours are converted to circadian time – 24 circadian hours per circadian day – permitting comparison between data from different free-running periods; M. Merrow *et al.*, unpublished). The timing of the peaks varies slightly from those presented in Lee *et al.*[2]. The subjective daytime (yellow background) is characterized by high levels of *frq* RNA, followed by FRQ protein synthesis. In the subjective night (gray background), WC-1 protein accumulates, which ultimately has positive effects on *frq* RNA accumulation, thereby completing the cycle. Some of the processes that occur through the circadian day are listed at the top of the graph, although precise timing of these events is not established for *Neurospora*. (b) and (c) *timeless* (*tim*) RNA and TIM protein levels in the *Drosophila* long-period *doubletime* (*dbt*) mutant, maintained in constant conditions (b) or in 12 h light:12 h dark cycles (c). Yellow background, actual lights on (c) or subjective daytime (the time when the lights would have come on) (b); gray background, the dark period in (c) and subjective night in (b). *Zeitgeber* time (c) is real hours from the initiation of the light (*zeitgeber*) incubation. Circadian time is a conversion from real hours representing one cycle in free-running conditions to a standard 24 h. (d) *frq* RNA levels in temperature cycles of different length. *Neurospora* was maintained in either 22-h cycles [11 h at 27°C (pink) followed by 11 h at 22°C (blue); $T = 22$] or 16-h cycles (8 h at 27°C, 8 h at 22°C; $T = 16$). *frq* RNA accumulates in different phases in these cycles: 'morning' in the long cycle and 'afternoon' in the shorter one (redrawn, with permission, from Ref. 18). Similarly to physiological experiments using whole organisms, as cycles decrease in length, the phase of the circadian output, a molecular readout shown here, lags[18,22]. In longer cycles, the phase is advanced.

Where the 24 hours come from

An intriguing aspect of this work is the 8-h lag between the appearance of FRQ and that of WC-1 in both the intact strain, and in the FRQ-knockout strain that was used for controlled expression of FRQ. This delay is one of a number of other processes that could slow the circadian feedback loop, including post-transcriptional mechanisms that control functional competence and stability of the proteins, such as subcellular localization, complex formation and phosphorylation[7] (Fig. 2a). How can we determine the delay potential in these processes in the molecular mechanism that structures a 24-h period? This is clearly problematic: how does one move beyond the correlative aspect of the lag time into causal inference?

One approach is indicated by experiments with *Drosophila* that compare profiles of molecular components in constant conditions with those during light–dark cycles[3]. The molecular circadian system in *Drosophila* is well described (see Fig. 1c for the transcription–translation feedback loop). PER and TIM dimerize and feed back negatively by inhibiting their activating dimer, CLOCK (CLK) and

CYCLE (CYC). An interconnected loop regulates expression of the activators themselves (through *clk* RNA)[12]. In addition, PER is increasingly phosphorylated over the course of the circadian day. In *Neurospora*, FRQ is also phosphorylated over the course of the day and disruption of this modification results in enhanced stability of the protein and correspondent lengthening of the circadian cycle[13]. Thus, post-transcriptional phosphorylation might be a common mechanism for control of transcription–translation loops. In *Drosophila*, *dbt* is thought to phosphorylate PER (*dbt* mutants are clock mutants with abnormal PER accumulation and phosphorylation patterns[14]), and thus DBT is crucial in the molecular circadian mechanism.

In addition to the use of constant conditions, many experiments are routinely performed using 24-h light:dark cycles (12 h light:12 h dark) with *Drosophila*. In both constant and cycling experimental conditions, there is a 4- to 6-h delay between the appearance of the RNA and the protein for the clock genes *per* and *tim*. This has also been noted in mouse for *mper1* and mPER1, and in *Neurospora* for *frq* and FRQ in constant conditions[7,15]. A similar lag between *per* and *tim* RNAs and their proteins is present when the long-period *dbt* mutant is held in constant darkness (Fig. 2b). However, when these flies are entrained in (synchronized with) a 24-h light cycle, the RNA and protein profiles are nearly superimposable (Fig. 2c). Thus, RNA and subsequent protein production are unlinked as sequential events. Underscoring the past decade of work demonstrating negative feedback within this loop, it is the declining phase of protein that determines the following RNA increase (i.e. disappearance of the protein allows RNA transcription to resume, see also Ref. 16). At least under this set of conditions, it appears that the segment of the cycle from protein decline to RNA accumulation represents an incompressible segment of the cycle.

Using circadian entrainment protocols to understand molecular functions

The difference in the RNA–protein relationship in constant and entrained conditions is a clue to the mechanisms regulating both RNA and protein production. However, we see that the two rhythms adopt different phase relationships depending on the experimental protocol. Such a phase change is also typical for core body temperature and activity rhythms in humans[17]. While under normal entrainment (i.e. living in the real world), our temperature rhythms reach their daily trough some hours before we wake, in constant conditions subjects initiate sleep at the temperature trough. Consequently, an established method to probe for a plasticity in the phase relationship of two rhythms is to use cycles of different lengths. All wild-type clocks have evolved under the selective pressure of a 24-hour day ($T = 24$) and have coordinated their physiology accordingly; that is, all rhythmic events have a defined relationship to the light–dark cycle. If they are, however, tested in entraining cycles of a different length (e.g. $T = 20$), the autonomy of the circadian system becomes apparent in the fact that the phase relationships of the circadian rhythms change (e.g. the rhythmic behavior appears later in relationship to the light–dark cycle). A caveat is that these protocols were designed and tested on whole organisms. Preliminary experiments, however, indicate that the molecular rhythms of clock components follow the same entrainment rules as were established for whole organisms (Fig. 2d)[18].

The *Drosophila dbt* mutant has a period of approximately 29 h in constant darkness, much longer than the wild-type rhythm of ~24 h. If the synchrony between the RNA and protein profiles in the *dbt* mutants in a 12 h:12 h light:dark cycle ($T = 24$) is due to plasticity of the phase relationship, one would predict reappearance of the characteristic 4- to 6-h lag if the mutants were held in T-cycles equal to their free-running period (i.e. $T = 29$). Furthermore, even wild-type flies could show synchrony of RNA and protein profiles if they were entrained by a cycle approximately 20% shorter than their free-running period (i.e. $T = 20$). If this were the case, the traditional view of RNA makes protein and, through some intermediates, protein inhibits transcription (as in Fig. 1a), would not suffice as the sole basis for circadian rhythms. One would then have to presume an additional oscillating process that impinges on transcription or on post-transcriptional processes.

Acknowledgements

The authors acknowledge the support of the Deutche Forschungsemeinshaft and the Friedrich-Bauer and Meyer-Struckmann Stiftungs, helpful comments from M. Mittag, R. Lucas, M. Rosbash and J. Loros; V. Suri and M. Rosbash for sharing data before publication; and invaluable assistance from C. Roenneberg.

References

1 Yan, O.Y. *et al.* (1998) Resonating circadian clocks enhance fitness in cyanobacteria. *Proc. Natl. Acad. Sci. U. S. A.* 95, 8660–8664
2 Lee, K. *et al.* (2000) Interconnected feedback loops in the *Neurospora* circadian system. *Science* 289, 107–110
3 Suri, V. *et al.* (2000) Two novel *doubletime* mutants alter circadian properties and eliminate the delay between RNA and protein in *Drosophila*. *J. Neurosci.* 20, 7547–7555
4 Hardin, P.E. *et al.* (1990) Feedback of the *Drosophila period* gene product on circadian cycling of its messenger RNA levels. *Nature* 343, 536–540
5 Elowitz, M.B. and Leibler, S. (2000) A synthetic oscillatory network of transcriptional regulators. *Nature* 403, 335–338
6 So, W.V. and Rosbash, M. (1997) Post-transcriptional regulation contributes to *Drosophila* clock gene mRNA cycling. *EMBO J.* 16, 7146–7155
7 Dunlap, J.C. (1999) Molecular bases for circadian clocks. *Cell* 96, 271–290
8 Dembinska, M.E. *et al.* (1997) Circadian cycling of PERIOD–β-galactosidase fusion protein in *Drosophila*: evidence for cyclical degeneration. *J. Biol. Rhythms* 12, 157–172
9 Liu, Y. *et al.* (1997) Thermally regulated translational control of *frq* mediates aspects of temperature responses in the *Neurospora* circadian clock. *Cell* 89, 477–486
10 Rosato, E. *et al.* (1997) Conceptual translation of timeless reveals alternative initiating methionines in *Drosophila*. *Nucleic Acids Res.* 25, 455–458
11 Crosthwaite, S.K. *et al.* (1997) *Neurospora wc-1* and *wc-2*: transcription, photoresponses, and the origin of circadian rhythmicity. *Science* 276, 763–769
12 Glossop, N.R.G. *et al.* (1999) Interlocked feedback loops within the *Drosophila* circadian oscillator. *Science* 286, 766–778
13 Liu, Y. *et al.* (2000) Phosphorylation of the *Neurospora* clock protein FREQUENCY determines its degradation rate and strongly influences the period length of the circadian clock. *Proc. Natl. Acad. Sci. U. S. A.* 97, 234–239
14 Price, J.L. *et al.* (1998) *double-time* is a novel *Drosophila* clock gene that regulates PERIOD protein accumulation. *Cell* 94, 83–95
15 Hastings, M.H. *et al.* (1999) Differential regulation of mPER1 and mTIM proteins in the mouse suprachiasmatic nuclei: new insights into a core clock mechanism. *J. Neurosci.* 19, 0:RC11 (1-7) (www.jneurosci.org)
16 Merrow, M. *et al.* (1997) Dissection of a circadian oscillation into discrete domains. *Proc. Natl. Acad. Sci. U. S. A.* 94, 3877–3882
17 Strogatz, S.H. (1987) Human sleep and circadian rhythms : a simple model based on two coupled oscillators. *J. Math. Biol.* 25, 327–347
18 Merrow, M. *et al.* (1999) Assignment of

circadian function for the *Neurospora* clock gene *frequency*. *Nature* 399, 584–586
19 Loros, J.J. and Feldman, J.F. (1986) Loss of temperature compensation of circadian period length in the *frq-9* mutant of *Neurospora crassa*. *J. Biol. Rhythms* 1, 187–198
20 Lakin-Thomas, P.L. (2000) Circadian rhythms: new functions for old clock genes? *Trends Genet.* 16, 135–142
21 Emery, P. *et al.* (2000) A unique circadian-rhythm photoreceptor. *Nature* 404, 456–457
22 Bruce, V. (1960) Environmental entrainment of circadian rhythms. *Cold Spring Harbor Symp. Quant. Biol.* 25, 29–48

M. Merrow*
T. Roenneberg
Institute for Medical Psychology,
Goethestrasse 31, 80336 Munich, Germany.
*e-mail: martha@imp.med.uni-muenchen.de

Genome Analysis

Gene conversions in genes encoding outer-membrane proteins in *H. pylori* and *C. pneumoniae*

I. King Jordan, Kira S. Makarova, Yuri I. Wolf and Eugene V. Koonin

***Helicobacter pylori* and *Chlamydia pneumoniae* are both pathogenic to humans. Their genomes have recently been completed, allowing detailed study of their evolution and organization. Here we describe an evolutionary analysis of the *H. pylori* and *C. pneumoniae* genes that encode their outer-membrane proteins. By comparing complete genome sequences of two *H. pylori* strains and two *C. pneumoniae* strains, we identify multiple independent conversions among these genes. Such recombination events might provide a selective advantage for these bacterial pathogens.**

H. pylori is a Gram-negative, human-specific gastric pathogen, which is a causative agent of chronic active gastritis as well as duodenal and gastric ulcers[1]. Chronic *H. pylori* infection can also have a role in the development of gastric carcinomas[2]. *Chlamydia pneumoniae* is another a human pathogen, which causes bronchitis and pneumonia[3]. In addition, *C. pneumoniae* infection has been associated with atherosclerosis[4]. The availability of complete genomic sequences of two *H. pylori* strains[5,6] and two *C. pneumoniae* strains[7,8] allows for detailed inferences concerning the genome organization and evolution of these medically important organisms to be made. We have employed these genomic sequence data in an evolutionary analysis of *H. pylori* and *C. pneumoniae* gene families that encode outer-membrane proteins.

Examination of the complete *H. pylori* genome sequences revealed the presence of the large Hop family of outer-membrane proteins[5,9]. All Hop-family members contain a conserved C-terminal domain. Members of the Hop family were initially characterized as porins with similar N-terminal amino acid sequences[10,11]. Subsequently, additional Hop-family members were found to be involved in adhesion to the gastric endothelium[12–14]. The two sequenced *C. pneumoniae* genomes also encode polymorphic families of outer-membrane proteins[8]. For example, the *C. pneumoniae* CWL029 genome encodes 21 members of the outer-membrane-protein family[7]. The biological role of this family is unknown, but the patterns of variation among the genes of the family indicate that molecular mechanisms exist to promote functional diversity of their encoded products.

Many of these outer-membrane proteins are probably important in pathogenesis and the presence of such proteins encoded by repetitive gene families indicates a possible role for the families in antigenic variation and host-defense evasion[15]. Several different mechanisms involving recombination among repeated genes can influence antigenic variation. Gene conversion is an intragenomic, nonreciprocal recombination event that results in identical (homogenized) gene sequences[16]. In bacterial pathogens, gene conversion is thought to be important in the generation of the repertoire of 'contingency genes' that mediate pathogen–host interactions[15]. In particular, there is evidence that antigenic variation in *Neisseria gonorrhoeae* pilus proteins is shaped by gene conversion between pilus genes[17]. In addition, recombination between *Mycoplasma genitalium* dispersed repetitive elements and the Mga operon probably generates antigenic variation in cellular adhesin proteins that are required for attachment of the organism to host epithelium[18]. Tomb *et al.*[5] hypothesized that similar recombination mechanisms could contribute to genetic, and subsequently antigenic, variation of the Hop gene family and its encoded products.

Although conversion has been invoked as an important mechanism of antigenic-variation maintenance, rigorously distinguishing this recombination mechanism from very recent intragenomic duplication is difficult. The complete genome sequences of two *H. pylori* and two *C. pneumoniae* strains provide the data necessary explicitly to test the hypothesis that conversion occurs between copies of gene family members that encode

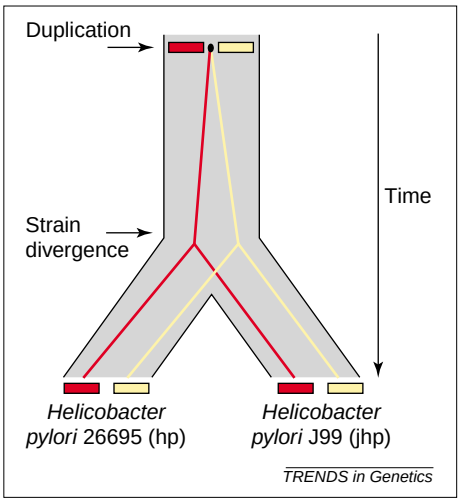

Fig. 1. Expected phylogenetic relationships among members of a gene family. Two strains, each containing two copies of a gene family, are represented. The strain lineage is shown in thick gray, and the gene lineages are shown with colored lines. Orthologs are indicated with the same color boxes and paralogs with different colors. Paralogous copies of a multi-gene family last shared a common ancestor at the time of gene duplication, whereas orthologous genes last shared a common ancestor at the time of strain divergence.

  PII: S0168-9525(00)02151-X